BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

As rescanning documents *will not* correct images, please do not report the images to the Image Problem Mailbox.

PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C12N 15/16, C07K 14/605 — A1

(11) International Publication Number: **WO 95/17510**

(43) International Publication Date: 29 June 1995 (29.06.95)

(21) International Application Number: PCT/DK94/00487

(22) International Filing Date: 22 December 1994 (22.12.94)

(30) Priority Data:
1440/93 23 December 1993 (23.12.93) DK

(71) Applicant *(for all designated States except US)*: NOVO NORDISK A/S [DK/DK]; Novo Allé, DK-2880 Bagsvaerd (DK).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: RASMUSSEN, Jesper, Skou [DK/DK]; Linde Allé 50, DK-2720 Vanløse (DK). THIM, Lars [DK/DK]; Skiftevej 22, DK-2820 Gentofte (DK). BJØRN, Søren, Erik [DK/DK]; Marie Grubbes Allé 47, DK-2800 Lyngby (DK).

(74) Common Representative: NOVO NORDISK A/S; Birgit Thalsø-Madsen, Corporate Patents, Novo Allé, DK-2880 Bagsvaerd (DK).

(81) Designated States: AM, AU, BB, BG, BR, BY, CA, CN, CZ, FI, GE, HU, JP, KG, KP, KR, KZ, LK, LT, LV, MD, MG, MN, NO, NZ, PL, RO, RU, SI, SK, TJ, TT, UA, US, UZ, VN, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG), ARIPO patent (KE, MW, SD, SZ).

Published
With international search report.

(54) Title: A METHOD OF PRODUCING GLUCAGON-LIKE PEPTIDE 1

(57) Abstract

Method of producing glucagon-like peptide 1 (GLP-1) 7-36 or an analogue thereof in a bacterium, the method comprising: (a) inserting in a suitable expression vector an expression cassette comprising a DNA construct encoding a GLP-1 precursor containing two or more consecutive DNA sequences coding for GLP-1 (7-36) or an analog thereof, the DNA construct encoding the GLP-1 precursor being preceded by a promoter sequence controlling the expression of the GLP-1 precursor, (b) transforming a suitable bacterium with the expression vector prepared in step (a), (c) culturing the transformed bacterium under suitable conditions permitting the expression of the DNA construct encoding the GLP-1 precursor, (d) recovering the resulting GLP-1 precursor from the bacterial culture, and (e) processing the GLP-1 precursor to GLP-1 (7-36) or an analog thereof.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	**GB**	United Kingdom	**MR**	Mauritania
AU	Australia	**GE**	Georgia	**MW**	Malawi
BB	Barbados	**GN**	Guinea	**NE**	Niger
BE	Belgium	**GR**	Greece	**NL**	Netherlands
BF	Burkina Faso	**HU**	Hungary	**NO**	Norway
BG	Bulgaria	**IE**	Ireland	**NZ**	New Zealand
BJ	Benin	**IT**	Italy	**PL**	Poland
BR	Brazil	**JP**	Japan	**PT**	Portugal
BY	Belarus	**KE**	Kenya	**RO**	Romania
CA	Canada	**KG**	Kyrgystan	**RU**	Russian Federation
CF	Central African Republic	**KP**	Democratic People's Republic of Korea	**SD**	Sudan
CG	Congo			**SE**	Sweden
CH	Switzerland	**KR**	Republic of Korea	**SI**	Slovenia
CI	Côte d'Ivoire	**KZ**	Kazakhstan	**SK**	Slovakia
CM	Cameroon	**LI**	Liechtenstein	**SN**	Senegal
CN	China	**LK**	Sri Lanka	**TD**	Chad
CS	Czechoslovakia	**LU**	Luxembourg	**TG**	Togo
CZ	Czech Republic	**LV**	Latvia	**TJ**	Tajikistan
DE	Germany	**MC**	Monaco	**TT**	Trinidad and Tobago
DK	Denmark	**MD**	Republic of Moldova	**UA**	Ukraine
ES	Spain	**MG**	Madagascar	**US**	United States of America
FI	Finland	**ML**	Mali	**UZ**	Uzbekistan
FR	France	**MN**	Mongolia	**VN**	Viet Nam
GA	Gabon				

A METHOD OF PRODUCING GLUCAGON-LIKE PEPTIDE 1

FIELD OF INVENTION

The present invention relates to a method of producing glucagon-like peptide or analogues or derivatives thereof.

BACKGROUND OF THE INVENTION

Mammalian proglucagon has been demonstrated to contain three different though homologous petides, glucagon, glucagon-like peptide 1 (also termed GLP-1) and glucagon-like peptide 2 (cf. L.C. Lopez et al., Proc. Natl. Acad. Sci. USA 80, 1983, pp. 5485-5489, and G.I. Bell et al, Nature 302, 1983, pp. 716-718). Prior to 1985, no definite biological activity of GLP-1 had been reported. However, in 1985 it was demonstrated that GLP-1(1-36)amide, like glucagon, stimulates insulin release from isolated precultured rat pancreatic islets in the presence of glucose in a dose-dependent manner (Schmidt, W.E. et al. Diabetologia 28 (1985) 704-7). This finding suggests that GLP-1(1-36)amide and related peptides might be useful in the treatment of type 2 diabetes. Due to its substantially closer sequence homology to glucagon and glucose dependent insulinotropic peptide, also referred to as GIP, Schmidt et al. suggested that an even stronger glucagon- and/or GIP-like biological activity could be expected with GLP-1(7-36) than with the intact peptide. In recent years, particular interest has focused on the GLP-1 fragments GLP-1(7-37) and GLP-1(7-36)amide and analogues and functional derivatives thereof. The amino acid sequence of GLP-1(7-36)amide and GLP-1(7-37) is given in formula I:

His-Ala-Glu-Gly-Thr-Phe-Thr-Ser-Asp-Val-Ser-Ser-Tyr-Leu-Glu-Gly-Gln-Ala-Ala-Lys-Glu-Phe-Ile-Ala-Trp-Leu-Val-Lys-Gly-Arg-X

(I)

which shows GLP-1(7-36)amide when X is NH_2 and GLP-1(7-37) when X is Gly-OH.

Thus, International Patent Application No. WO 87/06941 (to The General Hospital Corporation) relates to a peptide fragment which comprises GLP-1(7-37) and functional derivatives thereof and to its use as an insulinotropic agent. International Patent Application No. 90/11296 (to The General Hospital Corporation) relates to a peptide fragment which comprises GLP-1(7-36) and functional derivatives thereof and has an insulinotropic activity which exceeds the insulinotropic activity of GLP-1(1-36) or GLP-1(1-37) and to its use as an insulinotropic agent.

International Patent Application No. 91/11457 (to Buckley et al.) relates to effective analogues of the active GLP-1 peptides 7-34, 7-35, 7-36, and 7-37.

SUMMARY OF THE INVENTION

According to the present invention, it has been found advantageous to prepare various forms of GLP-1 by a process involving transformation of bacterial cells with DNA coding for GLP-1(7-36).

Accordingly, the present invention relates to a method of producing glucagon-like peptide 1 (GLP-1) 7-36 or an analogue thereof in a bacterium, the method comprising

(a) inserting in a suitable expression vector an expression cassette comprising a DNA construct encoding a GLP-1 precursor containing two or more consecutive DNA sequences coding for GLP-1(7-36) or an analogue thereof, the DNA construct encoding the GLP-1 precursor being preceded by a promoter sequence controlling the expression of the GLP-1 precursor,

(b) transforming a suitable bacterium with the expression vector prepared in step (a),

(c) culturing the transformed bacterium under suitable conditions permitting the expression of the DNA construct encoding the GLP-1 precursor,

(d) recovering the resulting GLP-1 precursor from the bacterial culture, and

(e) processing the GLP-1 precursor to GLP-1(7-36) or an analogue thereof.

In the present context, the term "analogue" is intended to indicate a peptide molecule which is functionally equivalent to GLP-1, i.e. which has substantially the same biological activity as native GLP-1, but whose amino acid sequence differs from that of native GLP-1 by one or more amino acids. Analogues of GLP-1 may suitably be prepared by modifying the DNA sequence coding for native GLP-1 by substituting one or more nucleotides in the sequence, or by inserting one or more codons into the sequence, adding one or more codons at either end of the sequence, or deleting one or more codons at either end of or within the sequence. Examples of suitable GLP-1 analogues are described in, e.g. WO 91/11457.

The term "expression cassette" is intended to indicate the DNA construct containing two or more consecutive DNA sequences coding for GLP-1(7-36) or analogue thereof preceded by a single promoter controlling the expression of the entire DNA construct. In other words, all DNA sequences coding for GLP-1(7-36) or analogues thereof present in the construct are expressed from this one promoter, rather than from separate promoters preceding each DNA sequence. Consequently, the DNA construct is expressed as one amino acid sequence (the GLP-1 precursor) containing two or more copies of GLP-1(7-36) or an analogue thereof.

In the work leading to the present invention, it was initially attempted to produce the peptide by introducing only one copy of the DNA sequence coding for GLP-1(7-36) into the expression vector. However, no recovery of the peptide was found on cultivation of bacterial cells transformed with the vector, either inside the cells or in the culture medium. On the other hand, when more than one copy of the DNA sequence coding for GLP-1(7-36) was inserted in tandem into the expression vector, it was surprisingly found that the transformed bacterial cells produced GLP-1 precursor in high yields (approximately 100 mg/l). The yield was not related to the copy number of the DNA sequence coding for GLP-1(7-36).

DETAILED DESCRIPTION OF THE INVENTION

The host cell used in the process of the invention may be any suitable bacterium which, on cultivation, produces large amounts of the GLP-1 precursor. Examples of suitable bacteria may be grampositive bacteria such as Bacillus subtilis, Bacillus licheniformis, Bacillus lentus, Bacillus brevis, Bacillus stearothermophilus, Bacillus alkalophilus, Bacillus amyloliquefaciens, Bacillus coagulans, Bacillus circulans, Bacillus lautus or Streptomyces lividans, or gramnegative bacteria such as *Escherichia coli*.It has been found that E.coli is able to produce a high yield of the GLP-1 precursor and is therefore a preferred host organism. In E. coli, the GLP-1 precursor is typically produced in the form of inclusion bodies. The transformation of the bacteria may for instance be effected by protoplast transformation or by using competent cells in a manner known per se (cf. Sambrook et al., Molecular Cloning. A Laboratory Manual, 2nd Ed., Cold Spring Harbor, NY, 1989).

According to the invention, the DNA construct encoding the GLP-1 precursor is preceded by a suitable promoter sequence, e.g. the promoter of the Bacillus stearothermophilus maltogenic amylase gene, Bacillus licheniformis α-amylase gene, Bacillus amyloliquefaciens BAN amylase gene, Bacillus subtilis alcaline protease gene, or Bacillus pumilus xylosidase gene, or by the phage Lambda P_R or P_L promoters, or the E. coli *lac* promoter. The DNA construct encoding the GLP-1 precursor may also be preceded by a ribosome binding site of of the Bacillus stearothermophilus maltogenic amylase gene, Bacillus licheniformis α-amylase gene, Bacillus amyloliquefaciens BAN amylase gene, Bacillus subtilis alcaline protease gene, Bacillus pumilus xylosidase gene, or E. coli *lac* gene.

The DNA construct of the invention comprising two or more consecutive DNA sequences encoding GLP-1(7-36) may suitably be prepared by ligating two or more cDNA sequences encoding GLP-1(7-36) which may, for instance, be obtained by preparing a mammalian, in particular human, cDNA library and screening for DNA sequences coding for GLP-1(7-36) by hybridization using synthetic oligonucleotide

probes in accordance with standard techniques (cf. Sambrook et al., Molecular Cloning: A Laboratory Manual, Cold Spring Harbor, 1982).

The DNA construct of the invention may also be produced by ligating two or more DNA sequences coding for GLP-1(7-36) or an analogue thereof prepared synthetically by established standard methods, e.g. the phosphoamidite method described by S.L. Beaucage and M.H. Caruthers, Tetrahedron Letters 22, 1981, pp. 1859-1869, or the method described by Matthes et al., EMBO Journal 3, 1984, pp. 801-805. According to the phosphoamidite method, oligonucleotides are synthesized, e.g. in an automatic DNA synthesizer, purified, ligated or alternatively assembled by PCR overlap extension as described in R.M. Horton, Gene 77, 1989, pp. 61-68, and cloned in an appropriate vector.

The DNA construct encoding the GLP-1 precursor may advantageously comprise three or more, in particular four or more or six or more, consecutive DNA sequences coding for GLP-1(7-36) or an analogue thereof.

When identical DNA sequences are placed in tandem, e.g. on an expression vector, they tend to rearrange by homologous recombination, i.e. part of the genetic information on the vector is lost as is the above-mentioned advantage of having multiple copies of the DNA sequence coding for GLP-1(7-36) or an analogue thereof present in the same DNA construct. To reduce the risk of rearrangement caused by homologous recombination, each of the DNA sequences coding for GLP-1(7-36) or an analogue thereof may be synthesized in such a way that it contains a high proportion of alternative codons (i.e. different codons specifying the same amino acid) to minimize the homology between the GLP-1(7-36) encoding sequences.

The expression vector comprising the DNA construct as described above may be any vector which is capable of replicating autonomously in a given host organism, typically a plasmid or bacteriophage. In the vector, the DNA sequence encoding the GLP-1 precursor should be operably connected to a suitable promoter sequence. The promoter may be any DNA sequence which shows transcriptional activity in the host

cell and may be derived from genes encoding proteins either homologous or heterologous to the host organism. Examples of suitable promoters are given above.

The vector may also comprise a selectable marker, e.g. a gene whose product confers antibiotic resistance, such as ampicillin, chloramphenicol or tetracyclin resistance, or the dal genes from B.subtilis or B.licheniformis.

The medium used to cultivate the cells may be any conventional medium suitable for growing bacteria. The GLP-1 precursor may be recovered from the medium by conventional procedures including separating the cells from the medium by centrifugation or filtration, if necessary after disruption of the cells to recover an intracellular product, precipitating the proteinaceous components of the supernatant or filtrate by means of a salt, e.g. ammonium sulphate, followed by purification by a variety of chromatographic procedures, e.g. ion exchange chromatography, affinity chromatography, or the like.

The GLP-1 precursor recovered in step (d) of the process of the invention is subsequently processed, in particular by enzymatic processing procedures known *per se* in the art. The enzyme used for processing should preferably be one which is specific for arginine residues, such as trypsin, trypsin-like protease derived from *Fusarium oxysporum* (WO 89/6270), clostripain (W.M. Mitchell et al., Methods Enzymol. 19, 1970, p. 635), mouse submaxillary gland protease (M. Levy et al., Methods Enzymol. 19, 1970, p. 672), or thrombin or other proteolytic enzymes of the blood coagulation cascade.

When expressed in *E. coli*, the N-terminal amino acid residue of the GLP-1 precursor will be a methionine. This implies that, after the GLP-1 precursor has been processed, a fraction of the recovered GLP-1(7-36) will be preceded by Met. To obtain authentic GLP-1(7-36), the N-terminal methionine will either have to be removed, e.g. by treatment with CNBr, or the fraction containing the Met-GLP-1(7-36) will have to be discarded. To reduce the costs of production and/or to improve the yield of authentic GLP-1(7-36), the DNA construct encoding the GLP-1 precursor is preferably preceded

by a presequence. In particular, the presequence codes for a prepeptide which has an arginine as the C-terminal amino acid residue. The prepeptide may suitably be cleaved off enzymatically, preferably with an arginine-specific protease, such as trypsin, trypsin-like protease derived from *Fusarium oxysporum*, clostripain, mouse submaxillary gland protease, or thrombin or other proteolytic enzymes of the blood coagulation cascade, so that it may be processed in the same step as the GLP-1 precursor. The prepeptide may subsequently be separated from the resulting GLP-1(7-36) by conventional procedures, such as chromatography, extraction and/or precipitation.

Likewise, the DNA construct encoding the GLP-1 precursor may comprise a spacer sequence between two or more of the GLP-1(7-36) encoding sequences. The spacer peptide encoded by the spacer sequence may comprise at least one, and preferably 2-5 amino acids. The identity of the amino acids in the spacer peptide is not critical, but it is assumed that the presence of acid amino acids in the spacer peptide would be an advantage as this would tend to expose the peptide for processing. The C-terminal amino acid of the spacer should preferably be arginine so that it may be cleaved off enzymatically by an arginine-specific enzyme as indicated above. Advantages of including one or more spacer sequences might be that the GLP-1 precursor will tend to precipitate in the form of inclusion bodies so as to provide a better yield of the precursor, and that processing of the precursor to GLP-1(7-36) may be optimized.

The DNA construct encoding the GLP-1 precursor may also be followed by a sequence coding for a post-peptide which, like the spacer peptide(s) serves the purpose of increasing the yield of the GLP-1 precursor, e.g. in the form of inclusion bodies. The number of amino acids in the post-peptide may suitably be from 1 to 10.

The GLP-1(7-36) or analogue thereof resulting from the processing of the GLP-1 precursor in step (e) of the method of the invention may be used *per se* as an active component in a pharmaceutical composition. Alternatively, it may be preferred to prepare a derivative of GLP-1(7-36) or analogue thereof. In the present context, the term "derivative" is intended to mean a peptidic compound resulting from any

enzymatic or chemical modification of GLP-1(7-36) resulting in a compound with insulinotrophic properties.

In a preferred embodiment, the GLP-1(7-36) may subsequently be converted to GLP-1(7-37) or GLP-1(7-36)amide. GLP-1(7-36) may suitably be converted to GLP-1(7-37) by C-terminal addition of Gly. Such conversion may, for instance, be effected by cleaving the C-terminal Arg36 from GLP-1(7-36) by means of a suitable enzyme, e.g. carboxypeptidase B or carboxypeptidase Y, in the presence of Arg-Gly dipeptide, resulting in replacement of the C-terminal Arg36 by Arg-Gly. Alternatively, it is contemplated that GLP-1(7-37) may be produced by processing of the GLP-1 precursor with a suitable enzyme capable of cleaving amino acid sequences at arginine residues (such as one of the enzymes indicated above) in the presence of Gly.

In an alternative process, GLP-1(7-37) may be prepared by constructing a DNA sequence encoding a GLP-1 precursor containing two or more DNA sequences encoding GLP-1(7-37) interspersed by sequence(s) coding for at least one arginine residue. The GLP-1 precursor may be prepared by the method of the invention as described above, and in step (e) of the method, the precursor may be processed with a suitable arginine-specific protease such as the mouse submaxillary gland protease and subsequently with carboxypeptidase B (an exopeptidase which cleaves off C-terminal arginine residues one at a time) so as to end processing at the C-terminal glycine.

GLP-1(7-36) may suitably be converted to GLP-1(7-36)amide by cleaving Arg36 from GLP-1(7-36) by means of a suitable enzyme, e.g. carboxypeptidase B or carboxypeptidase Y, in the presence of Arg-amide, resulting in the replacement of the C-terminal Arg36 by Arg-amide. The conversion may alternatively be carried out by cleaving the GLP-1 precursor with an Arg-specific enzyme in the presence of NH, in a medium which has a reduced water content, e.g. in an organic solvent.

Other possible derivatives of GLP-1(7-36) may be derivatives which are chemically modified at the C-terminal carboxylic acid group, e.g. pharmaceutically acceptable lower alkyl esters formed with the C-terminal carboxylic acid group, alkyl meaning e.g. methyl,

ethyl, propyl, isopropyl, butyl, or tert-butyl or alkylamide or dialkylamide wherein alkyl is as mentioned above.

The invention is further illustrated in the following examples which should not in any way be understood to limit the scope of the invention as claimed.

Example 1

Construction of a synthetic GLP1 cassette

A cassette of four GLP-1(7-36) coding units was assembled from synthetic oligonucleotides by polymerase chain reaction (PCR) overlap extension. In two separate PCR reactions single stranded oligonucleotide templates were joined. The extended products of these reactions were combined and joined in a third PCR reaction. PCR was performed as described in K.B. Mullis et al. Meth. Enzymol. 155 R. Wu (Ed.), 1987, pp. 335-350 using 10 μmol oligonucleotide as template, sequence specific oligonucleotide primers and AmpliTaq(TM) DNA polymerase (Perkin Elmer-Cetus). PCR was carried out for 25 cycles of amplification for 1 minute at 94°C, 1 minute at 55°C and 2 minutes at 72°C followed by 10 minutes at 72°C.

Oligonucleotides were synthesized on an ABI 394 DNA Synthesizer using phosphoamidite chemistry on a controlled pore glass support (Beaucage and Caruthers (1981) Tetrahedron Letters 22 pp.1859-1869).

The sequence of oligonucleotides used as templates was the following:

5'-ATTGGATCCCGTCACGCTGAAGGTACCTTCACCAGCGACGTTAGCAGCTAC
CTGGAAGGTCAGGCTGCTAAAGAATTCATCGCTTGGCTGGTTAAAGGTCG
TCATGCAGAGG-3'

5'-CTTCCGCGTGGCGGCCCTTTACGAGCCATGCAATAAACTCCTTTGCTGC
TTGGCCCTCGAGATACGACGATACATCAGAAGTAAAAGTGCCCTCTGCAT
GACGACCTTTA-3'

5'-AAAGGGCCGCCACGCGGAAGGGACGTTCACGAGTGACGTGAGTAGTTA
CTTGGAAGGGCAGGCGGCGAAAGAATTCATAGCGTGGTTGGTGAAAGGG
CGTCATGCCGAGG-3'

5'-ATATCTAGATTAGCGTCCCTTGACTAACCAGGCGATAAACTCCTTGGC
GGCTTGTCCCTCTAAATATGATGAGACATCTGATGTAAATGTTCCCTCGG
CATGACGCCCTTTC-3'

The sequence of the oligonucleotide primers was the following:

5'-ATTGGATCCCGTCACGCTGA-3'

5'-CTTCCGCGTGGCGGCCCTTT-3'

5'-AAAGGGCCGCCACGCGGAAG-3'

5'-ATATCTAGATTAGCGTCCCT-3'

The amplified PCR product was digested with restriction endonucleases BamHI and XbaI (New England Biolabs Inc., Berverly, Ma, USA) and subcloned into the cloning vector pSKII+ (available from Stratagene Cloning Systems, La Jolla, CA, USA) by the method described in (Sambrook et al. , Molecular Cloning: A Laboratory Manual 2nd Ed., Cold Spring Harbor, 1989). The nucleotide sequence of the PCR product was verified by the dideoxy chain termination metod of F. Sanger et al. PNAS, 74, 1977, pp. 5463-5467, using double stranded plasmid DNA as template and sequence specific oligonucleotides as sequencing primers.

The resulting double stranded PCR fragment had the following DNA sequence:

5'-ATTGGATCCC GTCACGCTGA AGGTACCTTC ACCAGCGACG TTAGCAGCTA
3-'TAACCTAGGG CAGTGCGACT TCCATGGAAG TGGTCGCTGC AATCGTCGAT

CCTGGAAGGT CAGGCTGCTA AAGAATTCAT CGCTTGGCTG GTTAAAGGTC
GGACCTTCCA GTCCGACGAT TTCTTAAGTA GCGAACCGAC CAATTTCCAG

GTCATGCAGA GGGCACTTTT ACTTCTGATG TATCGTCGTA TCTCGAGGGC
CAGTACGTCT CCCGTGAAAA TGAAGACTAC ATAGCAGCAT AGAGCTCCCG

CAAGCAGCAA AGGAGTTTAT TGCATGGCTC GTAAAGGGCC GCCACGCGGA
GTTCGTCGTT TCCTCAAATA ACGTACCGAG CATTTCCCGG CGGTGCGCCT

AGGGACGTTC ACGAGTGACG TGAGTAGTTA CTTGGAAGGG CAGGCGGCGA
TCCCTGCAAG TGCTCACTGC ACTCATCAAT GAACCTTCCC GTCCGCCGCT

AAGAATTCAT AGCGTGGTTG GTGAAAGGGC GTCATGCCGA GGGAACATTT
TTCTTAAGTA TCGCACCAAC CACTTTCCCG CAGTACGGCT CCCTTGTAAA

ACATCAGATG TCTCATCATA TTTAGAGGGA CAAGCCGCCA AGGAGTTTAT
TGTAGTCTAC AGAGTAGTAT AAATCTCCCT GTTCGGCGGT TCCTCAAATA

CGCCTGGTTA GTCAAGGGAC GCTAATCTAG ATAT-3'
GCGGACCAAT CAGTTCCCTG CGATTAGATC TATA-5'

Example 2

Inducible expression of GLP1 tetramer in E.coli

An oligonucleotide linker harboring sites for restriction endonucleases BamHI, XhoI and SpeI was introduced into the BamHI site of expression vector pET3a (available from AMS Biotechnology, UK, Ltd.). Oligonucleotides were synthesized on an ABI 394 DNA Synthesizer using phosphoamidite chemistry on a controlled pore glass support (Beaucage and Caruthers (1981) Tetrahedron Letters 22 pp.1859-1869).

The oligonucleotide linker had the following sequence:

5'-GATCCACTAGTCTCGAGA-3'
3'-GTGATCAGAGCTCTCTAG-5'

The GLP1 cassette was subcloned as a BamHI-XbaI fragment into the BamHI-SpeI site of the plasmid construct above by the method described in (Sambrook et al. , Molecular Cloning: A Laboratory Manual 2nd Ed., Cold Spring Harbor, 1989). The resulting expression cassette thus consists of four GLP1 (7-36) coding units preceded by sequences derived from bacteriophage T7 Gene 10 encoding a 15 residue leader peptide. The expression cassette translates into the following peptide sequence:

NH_2--
MASMTGGQQMGRGSRHAEGTFTSDVSSYLEGQAAKEFIAWLVKGRHAEGTF
TSDVSSYLEGQAAKEFIAWLVKGRHAEGTFTSDVSSYLEGQAAKEFIAWLVK
GRHAEGTFTSDVSSYLEGQAAKEFIAWLVKGR

The expression plasmid was introduced into E. coli BL21(DE3)/pLysS (available from AMS Biotechnology, UK, Ltd.) which was made competent according to procedures described in Sambrook et al., op.cit.

The plasmid harboring bacterial strain was incubated overnight at 37°C in with shaking in LB medium supplemented with ampicillin 100 ug/ml and chloramphenicol 30 ug/ml. The overnight culture was diluted 100 fold in fresh medium and icubated with shaking until OD600 = 0.8 - 1.0. Isopropyl (beta)-D-Thiogalactopyranoside (IPTG) was added to a concentration of 1 mM and incubation was continued. After induction for 3 -4 hours cells were harvested by centrifugation at 5.000 g for 10 minutes.

The bacterial pellet was frozen for 30 minutes at -20°C, then thawed and resuspended in 1/10 culture volumen of 50 mM Tris-Cl, pH=8.0, 2 mM EDTA, 0.1 % Triton X-100, causing the cells to lyse. Lysed cells were sonicated for 3 pulses for 30 seconds at medium output setting on a MSE soniprep 150 sonicator. The sonicate was centrifuged at 12.000 g for 15 min at 4°C. The pellet (inclusion bodies) was washed once in 1/20 culture volumen. Growth medium, lysate and inclusion bodies were kept for further analysis.

The preparation of inclusion bodies was solubilised in 4M guanidine hydrochloride and an aliquot was analysed by reverse phase HPLC. A Vydac 214TP54 column (C4, 4.6 x 250 mm) was used at a flow rate of 1.5 ml/min with 0.1% TFA as eluent A and 0.07% TFA in acetonitrile as eluent B. Equilibration was performed with 30% B and sample injection was followed by a linear gradient from 30-50% B in 40 min. The column effluent was monitored at 214 nm. Only one major peak was detected as seen in the chromatogram (Fig.1). The eluting material was collected and the N-terminal

amino acid sequence was found by Edman degradation on a Applied Biosystems 470A automatic sequencer, demonstrating that the collected material is pure and has the N-terminal of the GLP-1 precursor. The sequencer was stopped after 21 cycles and the obtained sequence covers 21 amino acids including the prepeptide and part of the first GLP-1 monomer. The result was confirmed by electrospray mass spectrometry analysis (SCIEX API *III*), that in addition demonstrate that the precursor consists of 4 units of GLP-1(7-36). The collected precursor was found have a molecular mass of 14541 Daltons. The theoretical calculated molecular weight of the entire precursor is 14677. The mass difference is due to a missing methionine residue at the N-terminal of the prepeptide as seen from the amino acid sequence analysis. The methionine residue has most probably been cleaved off by an enzymatic activity from the E.coli. However, the methionine residue is part of the peptide preceding the GLP-1 tetramer and will be cleaved off in the following conversion of the precursor to GLP-1 monomers.

Example 3

Constitutive expression of GLP1 tetramer in E. coli

A cassette of four GLP-1(7-36) coding units preceded by a methionine codon was made by performing PCR as described (K.B.Mullis op. cit.) using 100 ng of GLP1 teramer harboring plasmid, described in example 1 as template, sequence specific oligonucleotides as primers and AmpliTaq(TM) DNA polymerase (Perkin Elmer-Cetus). PCR was carried out for 25 cycles of amplification for 1 minute at 94°C, 1 minute at 55°C and 2 minutes at 72°C followed by 10 minutes at 72°C.

The resulting expression translates into the following peptide sequence:

NH_2–
MHAEGTFTSDVSSYLEGQAAKEFIAWLVKGRHAEGTFTSDVSSYLEGQAAK
EFIAW
LVKGRHAEGTFTSDVSSYLEGQAAKEFIAWLVKGRHAEGTFTSDVSSYLEG
QAAKEFIAWLVKGR

The sequence specific oligonucleotides used as primers were the following:

5'-ATTGCATCGATGCACGCTGAAGGTACC-3'

5'-ATATCTAGATTAGCGTCCCT-3'

Ends of the amplified PCR product were polished using DNA polymerase I large (Klenow) fragment (available from New England Biolabs Inc., Berverly, Ma, USA) as described (Sambrook et al. op. cit.). The PCR product was digested with restriction endonuclease ClaI and subcloned into a vector produced by digestion of pHD117-4SP13 (constructed substantially as described in EP 218 651) with restriction endonucleases ClaI and PvuII (available from New England Biolabs Inc., Berverly, Ma, USA). The nucleotide sequence of the PCR product was verified by the dideoxy chain termination metod (F. Sanger et al. op.cit.) using double stranded plasmid DNA as template and sequence specific oligonucleotides as sequencing primers.

The sequencing primers was the following:

5'-CTGTTGACAATTAATCATAGA-3'

5'-AAAGGGCCGCCACGCGGAAG-3'

The expression plasmid was introduced into E.coli MC1061 (T.V.Huynk et al., in DNA Cloning, vol. 1 (D.M.Glover, Ed.), IRL Press Ltd., Oxford, England, 1983 pp. 56-110) which was made competent according to procedures described in Sambrook et al. op.cit.

The plasmid harboring bacterial strain was incubated overnight at 37°C in with shaking in LB medium supplemented with ampicillin 100 ug/ml. The overnight culture was harvested by centrifugation at 5.000 g for 10 minutes. The bacterial pellet was frozen for 30 minutes at -20°C, then thawed and resuspended in 1/10 culture volumen of 50 mM Tris-Cl, pH=8.0, 2 mM EDTA, 100 ug/ml lysozyme (Sigma), 0.1 % Triton X-100.

Cells were incubated for 15 minutes at 37°C causing the cells to lyse. Lysed cells were sonicated for 3 pulses for 30 seconds at medium output setting on a MSE soniprep 150 sonicator. The sonicate was centrifuged at 12.000 g for 15 min at 4°C. The pellet (inclusion bodies) was washed once in 1/20 culture volumen. Growth medium, lysate and inclusion bodies were kept for further analysis.

Example 4

Conversion of the GLP-1 precursor containing four repeats by the use of Clostripain

Approx. 1 mg of the GLP-1 precursor was dissolved in 1 ml of 35 mM NaOH. The pH was adjusted to 2.2 by the addition of 1N HCl.

To 600 µl of the above precursor solution, 1800 µl was added of 0.1M Tris·HCl buffer containing 2.67 M urea, 2.5 mM DTT, 2.5 mM $CaCl_2$ and the pH was adjusted to 8.4. The enzymatic digestion was initiated by the addition of 1 µl clostripain solution (1.7 mg clostripain, Sigma C7403, Lot 43H6823 was dissolved in 1.7 ml of water), at an incubation temperature of 37°C. Samples (50 µl) were taken from the digestion mixture at t = 0 min, 2.5 min, 5 min, 7.5 min, 10 min, 15 min, 30 min, 1 h, 2 h and 24 h mixed with 50 µl of 0.2M HCl and analyzed by HPLC.

The samples were injected onto a Vydac 214TP54 column. The column was equilibrated with 30% (v/v) of MeCN in 0.1% (v/v) TFA. After the injection, the column was eluted with 30% (v/v) MeCN in 0.1% (v/v) TFA for 5 min, whereafter the concentration of MeCN was increased to 50% over 40 min. The flow was 1 ml/min, the temperature was 30°C and the absorbance at 214 nm was recorded.

In this system the undigested GLP-1 precursor eluted at a retention time of approx. 38.5 min whereas the GLP(7-36) peptide eluted at approx. 22.0 min.
The yield obtained as a function of time was as follows:

Table 1. Relative yield obtained by digestion of GLP-1 precursor with clostripain

Time	GLP-1 precursor	Intermediate products	GLP-1(7-36)
0 min	100%	0%	0%
2.5 min	100%	0%	0%
5.0 min	93%	7%	3%
7.5 min	80%	13%	7%
10 min	69%	19%	12%
15 min	49%	26%	25%
30 min	8%	30%	62%
1 h	2%	42%	56%
2 h	0%	73%	25%
24h	0%	100%	0%

As appears from table 1, the maximum yield of GLP-1(7-36) obtained by the digestion of the GLP-1 precursor with clostripain was 62% (incubation time: 30 min).

The identity of the GLP-1(7-36) peptide was confirmed by mass-spectrometry analysis as described in example 2. The GLP-1 peptide was isolated from the digest-mixture (30 min) and submitted to mass spectrometry analysis. The molecular weight determined was: 3299 ± 1 compared to the theoretical molecular weight of 3298.7.

The identity was further confirmed by N-terminal amino acid sequencing of the first 10 amino acid residues showing the expected sequence of His-Ala-Glu-Gly-Thr-Phe-Thr-Ser-Asp-Val-

Example 5

Conversion of the GLP-1 precursor containing four repeats by use of the trypsin-like protease from Fusarium oxysporum

Approx. 10 mg of GLP-1 precursor was dissolved in 2 ml of 35 mM NaOH at 4°C. Glycine was added to the solution to a resulting concentration of 35 mM and the pH was adjusted to 10.1.

The trypsin-like protease from Fusarium oxysporum was obtained from Novo Nordisk A/S (Product: SP387) in an aqueous solution containing 20 mg/ml.

The digestion was carried out at 37°C by the addition of 1.2 µl protease solution to the dissolved precursor. The mixture was analyzed by HPLC as described in example 4. The maximum conversion yield obtained after 5 min of digestion was 5.5%. This rather low yield was due to unspecific cleavages of the precursor at other residue than arginine, and can probably be further optimized by the use of a highly purified enzyme preparation.

Example 6

Conversion of the GLP-1 precursor containing four repeats by use of trypsin at high pH

Approx. 10 mg of GLP-1 precursor was dissolved in 2 ml 35 mM NaOH at 4°C. To 0.4 ml of this solution was added 0.15 ml of 35 mM glycine and the pH was adjusted to 11.5. The trypsin used was obtained from Novo Nordisk A/S (Porcine trypsin, batch P9284-30-4, 5200 USP/mg). The enzyme was dissolved in water in a concentration of 2 mg/ml. The digestion was carried out at 37°C by the addition of 10 µl trypsin solution to the dissolved precursor. Samples were withdrawn from the mixture at t = 0 min, 5 min, 15 min, 30 min, 1 h, 2 h and 3.5 h and analyzed in the HPLC system described in example 4.

The following results were obtained:

Table 2. Relative yield obtained by digestion of GLP-precursor with trypsin

Time	GLP-1 precursor	Intermediate products	GLP-1(7-36)
0 min	100%	0%	0%
5 min	90%	9%	1%
15 min	71%	25%	4%
30 min	59%	35%	7%
1 h	39%	52%	9%
2 h	17%	77%	6%
3½h	6%	88%	6%

The maximum yield was 9% and was obtained after 1 h of digestion.

Example 7

Conversion of the GLP-precursor containing four repeats by use of Endopeptidase Arg-C

The GLP-precursor was dissolved in a concentration of 5 mg/ml as described in example 5. The pH was adjusted to 9.0. Endopeptidase Arg-C (Boehringer, cat. no. 269590, lot no. 13170023-16) was dissolved in water to a concentration of 1.0 mg/ml.

The digestion was carried out at 37°C by addition of 15 µl enzyme solution to 300 µl of pressure solution. The digest mixture was analyzed by HPLC and described in example 4.

The maximum yield obtained was 2.7% after 21 h of digestion.

CLAIMS

1. A method of producing glucagon-like peptide 1 (GLP-1) 7-36 or an analogue thereof in a bacterium, the method comprising

(a) inserting in a suitable expression vector an expression cassette comprising a DNA construct encoding a GLP-1 precursor containing two or more consecutive DNA sequences coding for GLP-1(7-36) or an analogue thereof, the DNA construct encoding the GLP-1 precursor being preceded by a promoter sequence controlling the expression of the GLP-1 precursor,

(b) transforming a suitable bacterium with the expression vector prepared in step (a),

(c) culturing the transformed bacterium under suitable conditions permitting the expression of the DNA construct encoding the GLP-1 precursor,

(d) recovering the resulting GLP-1 precursor from the bacterial culture, and

(e) processing the GLP-1 precursor to GLP-1(7-36) or an analogue thereof.

2. A method according to claim 1, wherein the bacterium is *E. coli*.

3. A method according to claim 1 for producing GLP-1(7-36) derivatives.

4. A method according to claim 3, wherein the GLP-1(7-36) is subsequently converted to GLP-1(7-37) or GLP-1(7-36)amide.

5. A method according to claim 4, wherein GLP-1(7-36) is converted to GLP-1(7-37) by C-terminal addition of Gly.

6. A method according to claim 5, wherein Arg36 is cleaved from GLP-1(7-36) by means of a suitable enzyme, e.g. carboxypeptidase B or carboxypeptidase Y, in the

presence of Arg-Gly dipeptide, resulting in replacement of the C-terminal Arg36 by Arg-Gly.

7. A method according to claim 1, wherein GLP-1(7-36) is converted to GLP-1(7-37) by processing of the GLP-1 precursor with a suitable enzyme capable of cleaving amino acid sequences at arginine residues in the presence of Gly.

8. A method according to claim 4, wherein GLP-1(7-36) is converted to GLP-1(7-36)amide by cleaving Arg36 from GLP-1(7-36) by means of a suitable enzyme, e.g. carboxypeptidase B or carboxypeptidase Y, in the presence of Arg-amide, resulting in the replacement of the C-terminal Arg36 by Arg-amide.

9. A method according to claim 1, wherein the DNA construct encoding the GLP-1 precursor comprises three or more, in particular four or more consecutive DNA sequences coding for GLP-1(7-36) or an analogue thereof.

10. A method according to claim 9, wherein the DNA construct encoding the GLP-1 precursor comprises six or more consecutive DNA sequences coding for GLP-1(7-36) or an analogue thereof.

11. A method according to claim 1, wherein each of the DNA sequences coding for GLP-1(7-36) or an analogue thereof contains a high proportion of alternative codons to minimize the homology between the GLP-1(7-36) encoding sequences.

12. A method according to claim 1, wherein the DNA construct encoding the GLP-1 precursor is preceded by a presequence.

13. A method according to claim 12, wherein the presequence codes for a prepeptide which has an arginine as the C-terminal amino acid residue.

14. A method according to claim 12 or 13, wherein the presequence is cleaved off enzymatically, preferably with an arginine-specific protease.

15. A method according to claim 14, wherein the arginine-specific protease is trypsin, trypsin-like protease derived from *Fusarium oxysporum*, clostripain, mouse submaxillary gland protease, or thrombin or other proteolytic enzymes of the blood coagulation cascade.

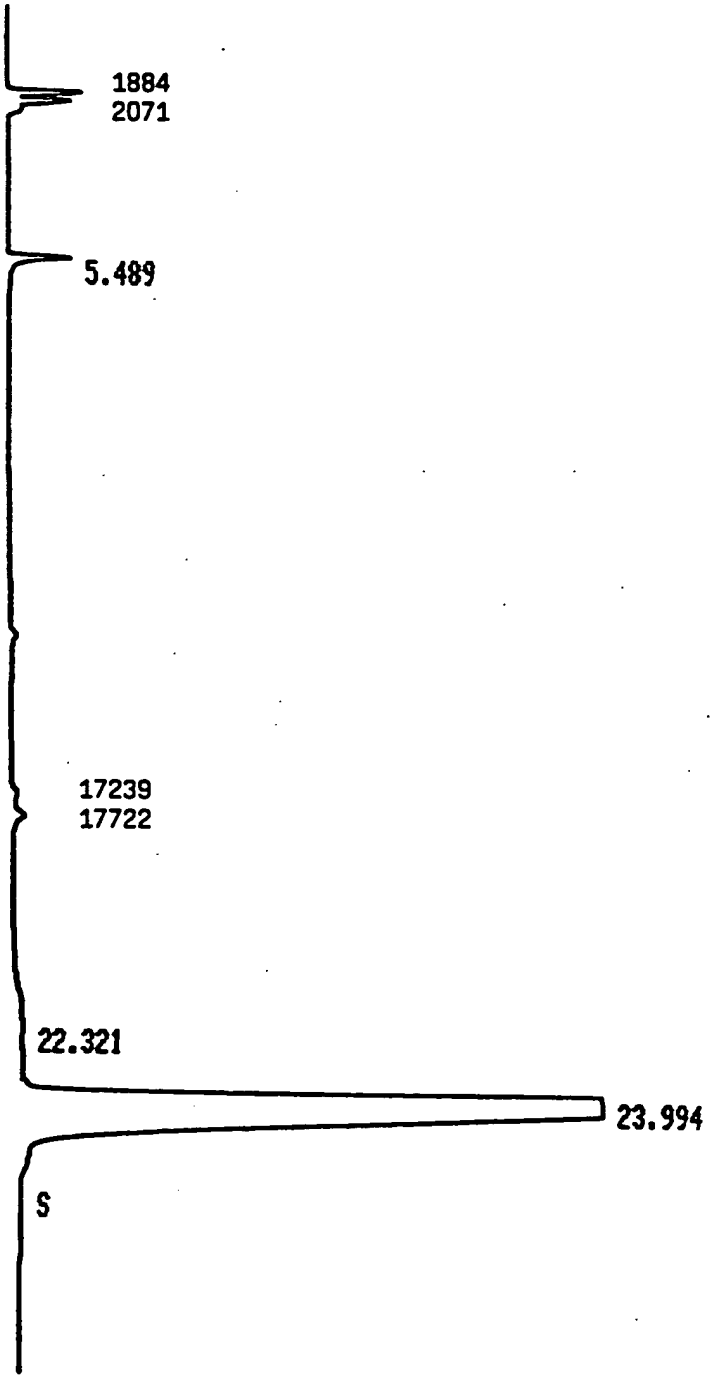

Fig. 1

INTERNATIONAL SEARCH REPORT

International application No.
PCT/DK 94/00487

A. CLASSIFICATION OF SUBJECT MATTER

IPC6: C12N 15/16, C07K 14/605
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC6: C07K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

SE,DK,FI,NO classes as above

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

WPI, MEDLINE, EMBASE

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US, A, 5202239 (S. JOSEPH TARNOWSKI ET AL), 13 April 1993 (13.04.93), column 2, line 13 - line 21; column 3 - column 8	1-3
Y	--	1-10,12-15
Y	US, A, 5118666 (JOEL F. HABENER), 2 June 1992 (02.06.92), see column 3, column 5, line 5 - line 14, column 6, lines 1-19	1-10,12-15
Y	--	
	WO, A1, 9001540 (CALIFORNIA BIOTECHNOLOGY INC.), 22 February 1990 (22.02.90), see page 24 - page 26, line 20, page 34, line 13 - line 29 and claims	1-10,12-15
	--	

[X] Further documents are listed in the continuation of Box C. [X] See patent family annex.

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"B" erlier document but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance: the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance: the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
23 March 1995	31 -03- 1995
Name and mailing address of the ISA/ Swedish Patent Offic Box 5055, S-102 42 STOCKHOLM Facsimile No. +46 8 666 02 86	Authorized officer Carl Olof Gustafsson Telephone No. +46 8 782 25 00

INTERNATIONAL SEARCH REPORT

Internati nal application No.
PCT/DK 94/00487

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US, A, 5120712 (JOEL HABENER), 9 June 1992 (09.06.92)	1-10,12-15
A	WO, A1, 9111457 (BUCKLEY, DOUGLAS), 8 August 1991 (08.08.91), see pages 4-7, 24-26 and claims.	1-8
A	US, A, 4826763 (KJELD NORRIS ET AL), 2 May 1989 (02.05.89)	1-15
P,X	US, A, 5322930 (S. JOSEPH TARNOWSKI ET AL), 21 June 1994 (21.06.94), see claim 6 and column 1 - column 3	1-3

INTERNATIONAL SEARCH REPORT

Information on patent family members

25/02/95

International application No.
PCT/DK 94/00487

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US-A- 5202239	13/04/93	CA-A- 2089094	08/02/92
		EP-A,A- 0542863	26/05/93
		JP-T- 6504036	12/05/94
		US-A- 5322930	21/06/94
		WO-A- 9202550	20/02/92
US-A- 5118666	02/06/92	AT-T- 110083	15/09/94
		DE-D,T- 3750402	01/12/94
		EP-A,B- 0305387	08/03/89
		SE-T3- 0305387	
		EP-A- 0587255	16/03/94
		JP-T- 1502746	21/09/89
		US-A- 5120712	09/06/92
		WO-A- 8706941	19/11/87
WO-A1- 9001540	22/02/90	AU-A- 4216689	05/03/90
		CA-A- 1320162	13/07/93
		EP-A,A- 0428615	29/05/91
		JP-T- 4501204	05/03/92
US-A- 5120712	09/06/92	AT-T- 110083	15/09/94
		DE-D,T- 3750402	01/12/94
		EP-A,B- 0305387	08/03/89
		SE-T3- 0305387	
		EP-A- 0587255	16/03/94
		JP-T- 1502746	21/09/89
		US-A- 5118666	02/06/92
		WO-A- 8706941	19/11/87
WO-A1- 9111457	08/08/91	EP-A,A- 0512042	11/11/92
US-A- 4826763	02/05/89	AU-B- 587218	10/08/89
		AU-A- 5247286	31/07/86
		CA-A- 1294232	14/01/92
		EP-A,B- 0189998	06/08/86
		SE-T3- 0189998	
		IE-B- 58726	03/11/93
		JP-A- 6292594	21/10/94
		JP-A- 61181398	14/08/86
US-A- 5322930	21/06/94	CA-A- 2089094	08/02/92
		EP-A,A- 0542863	26/05/93
		JP-T- 6504036	12/05/94
		US-A- 5202239	13/04/93
		WO-A- 9202550	20/02/92